Automatic Detection of Clear-Sky Periods From Irradiance Data

Benjamin H. Ellis, Michael Deceglie, and Anubhav Jain

***Abstract*—Recent degradation studies have highlighted the importance of considering cloud cover when calculating degradation rates, finding more reliable values when the data are restricted to clear sky periods. Several automated methods of determining clear sky periods have been previously developed, but parameterizing and testing the models has been difficult. In this paper, we use clear sky classifications determined from satellite data to develop an algorithm that determines clear sky periods using only measured irradiance values and modeled clear sky irradiance as inputs. This method is tested on global horizontal irradiance (GHI) data from ground collectors at six sites across the United States and compared against independent satellite-based classifications. First, 30 separate models were optimized on each individual site at GHI data intervals of 1, 5, 10, 15, and 30 min (sampled on the first minute of the interval). The models had an average $F_{0.5}$ score of 0.949 $\pm$ 0.035 on a holdout test set. Next, optimizations were performed by aggregating data from different locations at the same interval, yielding one model per data interval. This paper yielded an average $F_{0.5}$ of 0.946 $\pm$ 0.037. A final, "universal" optimization that was trained on data from all sites at all intervals provided an $F_{0.5}$ score of 0.943 $\pm$ 0.040. The optimizations all provide improvements on a prior, unoptimized clear sky detection algorithm that produces $F_{0.5}$ scores that average to 0.903 $\pm$ 0.067. Our paper indicates that a single algorithm can accurately classify clear sky periods across locations and data sampling intervals.**

I. Introduction

SOLAR resource availability and variability have become active topics of study given the accelerating adoption of photovoltaic (PV) technologies. By accounting for solar irradiance variation in PV monitoring and degradation analyses, one can isolate outages, anomalies, and calculate degradation under consistent solar resource availability. For example, recent work from the National Renewable Energy Laboratory (NREL) and SunPower Corporation demonstrated that introducing clear sky modeling and filtering into retrospective degradation rate analyses of PV system data can substantially reduce uncertainty and bias in the extracted degradation rate due to, for example, drifting irradiance sensors [1]. In their work, skies were determined to be clear if the measured global horizontal irradiance (GHI) was within 20% of modeled clear-sky GHI. The authors acknowledge that a static threshold is not an ideal solution given that they observed that irradiance sensors can drift by up to 2%/year—highlighting the need for more accurate methods to detect clear-sky periods.

In this paper, clear-sky periods will be defined as periods throughout a given day that have measured GHI values indistinguishable from modeled clear-sky GHI. This definition is the same as the previous work's definition, as well as other recent research efforts [2], [3]. This selection is made for two reasons. First, many PV sites do not track meteorological data related to cloud cover. Relying on such data would limit the applicability of the proposed methodology. This paper will specifically focus on the comparison of measured GHI to modeled GHI (although other measurements, such as direct normal irradiance or diffuse horizontal irradiance, could in principle be used). We note that within this definition, a clear sky period may still include the presence of clouds, aerosols, or other particulates so long as those conditions do not significantly affect the irradiance that PV sites collect. The task in this paper is to best formalize this definition into a working algorithm for clear sky detection.

Classifying sky conditions using ground-based irradiance measurements has been previously researched. Many researchers rely on the clearness index (the ratio of GHI observed from ground measurements to the corresponding extraterrestrial irradiance) to determine if skies are clear, cloudy, or even determine what types of clouds are present at a given time [4]–[8]. None of these studies agreed on thresholds for the clearness index to determine clear sky periods; the minimum threshold ranged from 0.6 to 0.7. Similar work has also considered the diffuse fraction of GHI in determining clear skies [9] and correlated meteorological data with the clearness index and diffuse fraction data [10]. Progress has also been made in determining cloud cover and cloud types from satellite data [11], [12]. Advancements in computer vision and image recognition have also resulted in methods to derive cloud conditions from ground-based images of the sky [13]–[15]. More recently, a method to determine sky clarity by extracting five features from measured GHI curves and comparing them with features extracted from expected clear sky irradiance has been developed [3]. Extending

This work was primarily supported as part of the Durable Modules Consortium, an Energy Materials Network Con-sortium funded by the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Solar Energy Technologies Office. This work was also authored in part by the Alliance for Sustainable Energy, LLC, the Manager and Operator of the National Renewable Energy Laboratory for the U.S. Depart-ment of Energy, under Contract No. DE-AC36-08GO28308. The work at the Lawrence Berkeley National Laboratory was supported under Contract No. DE-AC02-05CH11231. *(Corresponding author: Anubhav Jain.)*

B. H. Ellis and A. Jain are with the Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA (e-mail: bhellis@lbl.gov; ajain@lbl.gov).

M. Deceglie is with the National Renewable Energy Laboratory, Golden, CO 80401 USA (e-mail: michael.deceglie@nrel.gov).

this most recent work, the goal of our paper is to predict clear sky periods automatically (across locations and data sampling intervals) using only GHI data as an input.

II. Methods

This paper will demonstrate a method to automatically determine thresholds for existing clear sky classification algorithms by combining available ground-based irradiance data with satellite data. In this paper, we focus solely on the detection algorithm developed previously by Reno and Hansen because it only requires GHI as an input [3]. To distinguish clear and cloudy skies, this method (subsequently called the "PVLIB method") calculates five features for discrete windows of time along time-series GHI data. Four of the five features are calculated for both measured GHI and modeled clear-sky GHI ($\mathrm{GHI_{CS}}$). The difference between GHI and $\mathrm{GHI_{CS}}$ of these four features must be within some tolerance to label the period as clear sky, i.e., these features determine whether modeled clear-sky GHI and measured GHI are in agreement. The final feature, standard deviation of the slopes, must be below some absolute tolerance value and ensures that the GHI curve is smooth rather than irregular. We will describe the five features developed for the PVLIB method concisely—an in-depth justification of each is described in the original work. We represent measured irradiance as *G* and clear-sky irradiance as G_{CS}. The modeled clear sky was obtained using the Ineichen model via the PVLIB-Python implementation [16], [17]. Other clear sky models can be used with this method, although we did not test them in this study.

In the following equations, *i* will represent the index of single data point (GHI measurement) in a window that contains *n* elements. The first two features described by Reno and Hansen are the differences between the average and maximum measured and modeled irradiance over a given window of time [3]

$$x_1 = \left|\overline{G} - \overline{G}_{\mathrm{CS}}\right| \tag{1}$$

$$x_2 = \left|\max\left(\{G_i, \ldots, G_n\}\right) - \max\left(\{G_{\mathrm{CS},i}, \ldots, G_{\mathrm{CS},n}\}\right)\right|. \tag{2}$$

The third feature is the difference in line length of the irradiance-versus-time curves

$$x_3 = G^{\mathrm{LL}} \; - \; G^{\mathrm{LL}}_{\mathrm{CS}} \tag{3}$$

in which the line length is calculated as follows:

$$G^{\mathrm{LL}} = \sum_{i=1}^{n-1} \sqrt{\left(G_{i+1} - G_i\right)^2 + \left(t_{i+1} - t_i\right)^2}. \tag{4}$$

The fourth feature is the standard deviation of the slope between adjacent data points in a time window normalized by the average irradiance

$$x_4 = \frac{1}{\overline{G}} \sqrt{\frac{1}{n-1} \sum_{i=1}^{n-1} \left(s_i - \bar{s}\right)^2} \tag{5}$$

where s_i values are the slopes between adjacent points and $\bar{s}$ is their average. The fifth feature is the maximum difference between the slope of the measured and modeled irradiance over a window

$$x_5 = \max\left(\left|\alpha_G - \alpha_{G_{\mathrm{CS}}}\right|\right) \tag{6}$$

$$\alpha = \frac{G_{i+1} - G_i}{t_{i+1} - t_i}. \tag{7}$$

We note that in the original work, time is not included in the denominator of the slopes because that work implicitly assumes 1-min data [3]. Our paper will generalize many data intervals and thus benefits from this normalization.

In the original work, the authors demonstrated the model's efficacy by determining thresholds for the five features manually. Results from the model were visually inspected and compared with previous clear sky models and satellite data. Based on these results, the PVLIB method provides default thresholds, assuming that irradiance data are measured every minute, for each of the features presented. These parameterizations include: a window size of 10 min, an absolute threshold of 75 W/m^2 for average and maximum difference of GHI and $\mathrm{GHI_{CS}}$, a line length difference from –5 to 10, a threshold for the standard deviation of differences of 0.005, and a maximum absolute difference between adjacent points of 8 W/m^2. The thresholds listed work well for minutely data intervals; however, as we will demonstrate in the results section, they do not generalize well to higher interval data. We note that these features reflect the definitions from the original article [3]. We note that the implementation in version 0.5.2 of PVLIB-Python (the version as of the time this work was initiated) contains inconsistencies between the publication and implementation that were fixed for this paper—the source code used in this paper can be found in PVLIB-Python version 0.5.2+13.g3f9dc9d.

In cases where the default PVLIB parameters are inadequate, researchers must optimize these parameters individually. This procedure adds additional time, complication, and nonstandardization (variability) to monitoring and studying PV systems—especially large studies encompassing multiple sites. This motivated our paper to automatically tune PVLIB parameters by scoring PVLIB sky classifications versus known clear-sky periods using a combination of ground and satellite data.

In this paper, we correlate ground-based irradiance data with satellite-derived cloud coverage to automatically determine optimal thresholds for the PVLIB detection scheme at multiple intervals. The ground collectors are all members of the NREL Measurement and Instrumentation Data Center (MIDC) network. Stations in this network monitor solar radiation and meteorological data across the United States. For this work, we have selected six different locations for study [18]–[23]. Their information is listed in Table I. These sites were selected because they each collected GHI at 1-min intervals and have at least four years of overlap with satellite data in the National Solar Radiation Database (NSRDB). The ground-based irradiance measurements serve as inputs to the PVLIB clear sky classifier. The span and amount of data for each MIDC site is detailed in Table I. Note that these dates are the total span of data. When reporting our results, the final calendar year of each data set was held out for testing and scoring (i.e., not used to optimize PVLIB model parameters).

TABLE I
MIDC STATION METADATA

Station ID	Elevation (meters)	Latitude (degrees)	Longitude (degrees)	Begin date	End date	Number of ground-based data points	Number of NSRDB data points
HSU	36	40.88	-124.08	05-02-2007	01-01-2014	3507840	116832
LRSS	1860	39.61	-104.58	06-01-2008	01-01-2014	2937600	97872
ORNL	245	35.93	-84.31	01-01-2010	01-01-2016	3155040	105120
LMU	27	33.97	-118.42	05-01-2010	01-01-2016	2982240	99360
UNLV	615	36.11	-115.14	01-01-2007	01-01-2014	3682080	122640
BMS	1829	39.74	-105.18	09-01-2007	06-01-2012	2498400	83280

Begin and end dates are in MM-DD-YYYY format.

Satellite data has been collected from the NSRDB Physical Solar Model version 2 [12]. This resource provides cloud cover data at 30-min intervals across the United States at a spatial resolution of 4×4 km. The NSRDB is capable of assigning several cloud classifications, although a binary classifier to determine "clear" or "not clear" was developed for the application in mind. Clear periods were determined by the "clear" NSRDB determination, with the constraint that expected clear-sky irradiance must be greater than 0 W/m^2 to remove clear periods from dusk to dawn. All other classifications were determined to be cloudy. Periods with missing cloud types were not used for training or scoring. Only points occurring at every 30th min were used for training and scoring—a limitation necessitated by the satellite collection interval. All lower interval data points measured from the ground were used for calculating the five model features.

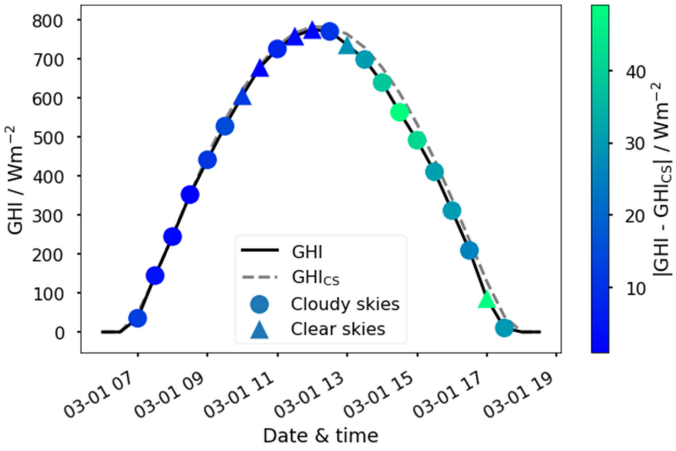

Fig. 1. Difference between GHI and GHI$_{CS}$ throughout a sample day indicates potential inconsistencies in satellite-based clear sky labels and GHI data.

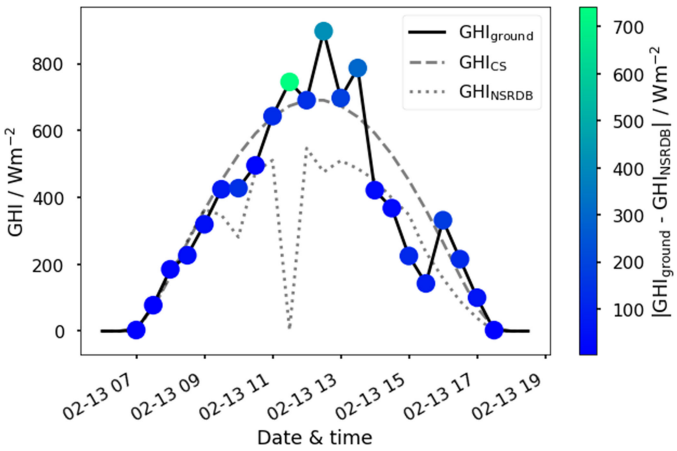

Fig. 2. Differences between GHI measured from ground-based collectors and NSRDB for a sample day indicates discrepancies in satellite-based GHI data and ground-based GHI data.

III. CALCULATION

Before parameter optimization, several data preprocessing steps removed erroneous data points. First, periods missing cloud data from NSRDB data sets were removed. Second, periods from dusk to dawn were also removed from scoring (expected clear sky irradiance had to be greater than 0 Wm^{-2}). In the third cleaning step, points having GHI measurements in very strong agreement with modeled clear-sky GHI values (difference less than 50 W/m^2 averaged over 1 h) but labeled as "cloudy" by NSRDB determinations were removed from training and testing. This issue is illustrated in Fig. 1, where all "cloudy" points as classified by NSRDB (circles) have an absolute difference of less than 50 Wm^{-2} from the modeled clear-sky GHI. In the same figure, we see that clear points (triangles) also have a difference of less than 50 Wm^{-2}. Removing these "conflicting" points, in which there is clear a disagreement between irradiance data and NSRDB cloud classification, will help the models more reliably distinguish clear and cloudy skies. A similar cleaning step removed periods labeled as clear but that had an average absolute difference between GHI and GHI$_{CS}$ (over 1 h) greater than 50 Wm^{-2}. The next preprocessing step removed data points from training and scoring where GHI values from NSRDB and ground measurements differed by a large degree. This step ensured that GHI from satellite data agreed with irradiance directly measured at ground locations. Discrepancies between satellite and ground GHI measurements are presented in Fig. 2. Points where the average difference between ground and NSRDB GHI was greater than 50 Wm^{-2} over 1 h were filtered out of training, testing, and scoring. These preprocessing steps guaranteed that an intuitive and consistent definition of clear skies (measured GHI is indistinguishable from modeled GHI$_{CS}$) was reflected in the data and helped compensate for inaccuracies in the training data set.

In this paper, clear-sky threshold parameters for the five features were determined using data from each of the 6 MIDC sites at intervals of 1, 5, 10, 15, and 30 min individually, resulting in 30 different optimizations. Data from the MIDC sites were collected at a 1-min interval and were downsampled with no

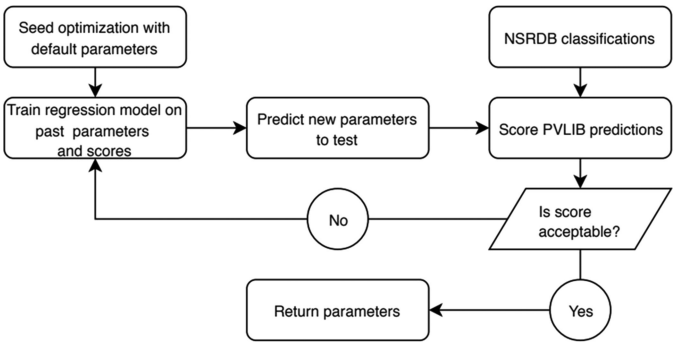

Fig. 3. Flowchart of the parameter optimization procedure.

smoothing for the 5, 10, 15, and 30 min optimizations. In this paper, we define downsampling as selecting every *i*th minute from the data set (e.g., for the 10-min interval optimization, we selected every 10th min from the original, 1-min interval data). A further discussion of this topic is in the supporting information.

The performance of each model is quantified using the F_β score

$$F_\beta = (1+\beta^2) \times \frac{\text{precision} \times \text{recall}}{(\beta^2 \times \text{precision}) + \text{recall}}. \quad (8)$$

The F_β score is a weighted mean between the precision and recall. In this paper, β was set to 0.5, giving greater weight to precision. This metric will minimize the occurrence of false positives (i.e., cloudy periods incorrectly classified as clear), which is of interest for the application in mind.

The optimal parameters for our model were determined by maximizing the $F_{0.5}$ scores between PVLIB predictions and satellite classifications in the training data set, as illustrated in Fig. 3. The optimization algorithm iteratively searches for points in the PVLIB parameter space. Before a new set of parameters is tested, a gradient-boosted forest regression is trained on previous PVLIB parameters and $F_{0.5}$ scores. The new sets of PVLIB parameters are picked from regions, as predicted by the regression model, where the $F_{0.5}$ score is expected to increase. The selected parameters are scored and used in future iterations to train the regression model. Each optimization procedure was terminated after testing 250 sets of PVLIB parameters.

IV. Results and Discussion

Three sets of results are presented that illustrate the $F_{0.5}$ scores of various newly trained PVLIB parameters. These scores reflect the degree of agreement between classifications from our algorithm and satellite data used in conjunction with the NSRDB Physical Solar Model version 2 [12]. In method A, we individually optimize a unique set of PVLIB parameters for each combination of location and data interval. In method B, we aggregate the data by interval across sites to produce a different set of optimized PVLIB parameters tailored to each data interval. Finally, in method C, we consider only a single "universal" parameter set for the entire data set across all data intervals and all sites across the MIDC network. All $F_{0.5}$ scores presented in this paper were calculated using a hold-out, test

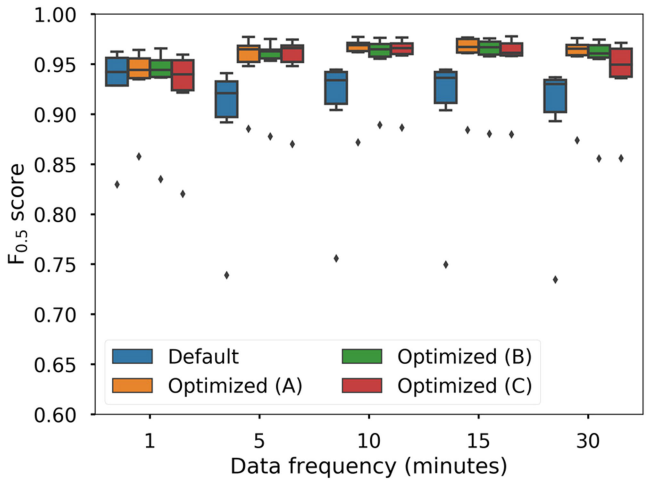

Fig. 4. Boxplots representing $F_{0.5}$ scores averaged across all six MIDC locations at each interval using default and optimization PVLIB thresholds. (a) Parameters were optimized individually at each site and interval. (b) Parameters were optimized by combining data at each interval. (c) Parameters are the result of optimizing on all site and intervals simultaneously. All scores were calculated on the same hold-out data set. Black points indicate that they are below the 1.5 × IQR (interquartile range).

year that was not observed during the parameter optimization process.

$F_{0.5}$ scores for default and optimized PVLIB parameters are summarized in Fig. 4. The default scores were all computed using default (previously published [3]) PVLIB parameters. For default scores, the window size was set to the PVLIB default of 10 min for minutely data and to 1-h window size for lower interval data (i.e., 5, 10, 15, and 30-min intervals). Optimized PVLIB parameters and scores were calculated using a fixed 1-h window for all intervals.

Fig. 4 illustrates that default parameters compare well with optimized parameters on minutely data. This is an expected result because the original PVLIB parameters were tuned for this data interval. This figure also demonstrates that as data interval increases, the performance of the default parameters becomes worse. We note again that the data interval was increased (e.g., from 1-min to 5-min) by selecting every *i*th minute from the data set; no smoothing was performed. We see that the optimized scores are consistent from minute data to 30-min data using all three optimization approaches.

Detailed scores for each of the parameter sets is presented in Fig. 5. Again, our results indicate that there is very little difference between optimized parameters and default parameters for minute data. As the interval increases, we see that scores from default parameter sets and optimized parameter sets diverge. In both default and optimized parameter sets, we see that Oak Ridge National Laboratory (ORNL) performs significantly worse than other sites (it also represents the outlier in Fig. 4). This is true even for method A, where that site is independently optimized, indicating a fundamental challenge with improving performance at that site rather than some characteristic difference from the other sites. We believe the poor performance of ORNL stems from disagreement between satellite and ground

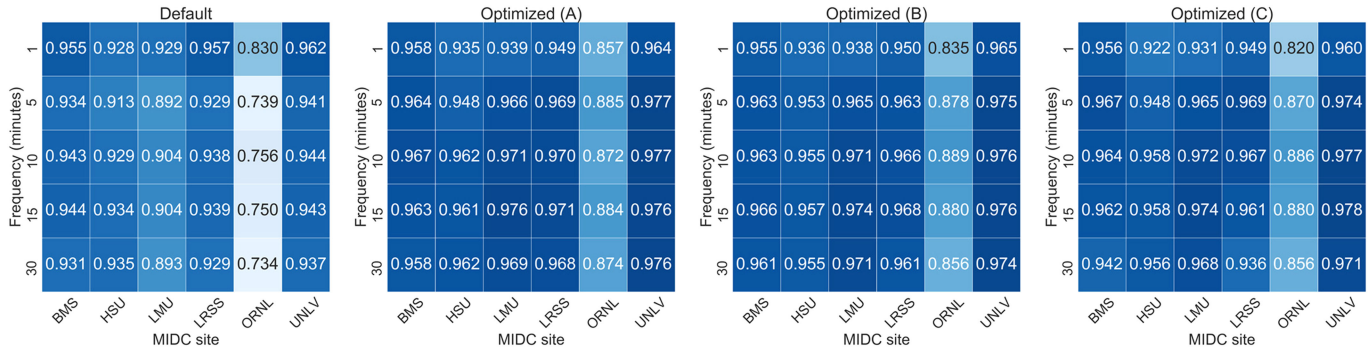

Fig. 5. Detailed $F_{0.5}$ scores using default and optimized parameters. (a) Parameters were optimized individually at each site and interval (frequency). (b) Parameters were optimized by combining data at each interval. (c) parameters are the result of optimizing on all site and intervals simultaneously. All scores were calculated on the same hold-out data set.

measurements of GHI at that location (see supporting information). Despite the poor performance of ORNL data relative to other sites, it is encouraging to see that the optimization scores still outperform PVLIB's default parameters on longer interval data.

The set of three optimizations was performed to demonstrate that optimal thresholds for the PVLIB detection scheme can be considered to be location and interval agnostic (Optimized (C) results in Figs. 4 and 5). Optimization (A) results indicate the performance to expect when PVLIB parameters are optimized at a specific location and for a specific interval. These parameters have an $F_{0.5}$ score equal to $0.949 \pm .035$ (The mean $F_{0.5}$ score plus/minus the standard deviation). Although this approach would yield the highest possible score for any given data set, it would be unfeasible to perform this optimization in a large-scale degradation study that spans tens or hundreds of locations. Optimization (B) parameters were determined by training on data aggregated at each interval. These results reveal that PVLIB parameters are not unique to location and provide performance similar to parameters optimized specific to a location and interval. On average, these parameters provided an $F_{0.5}$ score of 0.946 ± 0.037. This encouraging result motivated the question of whether there is a set of "universal" parameters that provide good $F_{0.5}$ score regardless of site location or data interval. Optimization (C) parameters present evidence that a single set of PVLIB parameters work across locations and intervals for less than 1 h. Our findings demonstrate that PVLIB parameters obtained by performing the optimization on all locations and intervals simultaneously provide nearly identical performance to parameters optimized for specific locations and data intervals. The universally optimized parameters produced an $F_{0.5}$ score of 0.943 ± 0.040.

We are optimistic that the parameters from optimization (C) are in fact universal because they perform nearly as well the single-site optimization, suggesting that the algorithm is not too sensitive to the specifics of any single site. In addition, because the Reno and Hansen method relies on physically relevant characteristics of a clear-sky GHI profile, we expect the application at new sites to be robust. However, we have not validated the method globally, and care should be taken when using these results for new sites outside this study.

TABLE II
DEFAULT AND OPTIMIZED PVLIB PARAMETERS (FOR ALL SITES AND DATA INTERVALS) WHEN USING A 60-MIN TIME WINDOW TO CALCULATE FEATURES

	Default	Optimized
Mean difference	75	79.144
Max. difference	75	59.152
Lower line length	-5	-41.416
Upper line length	10	77.789
Std. dev. of slopes	0.005	0.00745
Max diff. of slopes	8	68.579

Optimal parameters, applicable to all sites and intervals, are compared with default parameters in Table II. It is important to note that the default parameters were determined manually using a 10-min window to calculate features [3]. The optimization procedure in this paper used 60-min windows to calculate features. We chose the 60-min window because it is applicable to all data intervals used in the study. It is interesting to note that the optimal thresholds for mean and max difference and the standard deviation of slopes are similar to those proposed in the original work. However, Table II indicates that the remaining thresholds are quite different from the defaults. We suggest that the differences in the thresholds are due to our optimization for higher interval data in contrast with the 1-min data used to optimize the default thresholds [3].

To visually demonstrate the effect of the revised parameters, a sample of the performance of the universal, optimized PVLIB parameters is illustrated in Fig. 6. Results are plotted for the Baseline Measurement System (BMS) site as it is representative of the performance across most sites. Upon visual inspection, it appears that default parameters are more conservative than the optimized parameters; in general, the optimized parameters label more periods as clear. In some cases, the optimized parameters may perhaps mislabel cloudy periods (morning of 2012-04-08, 15-min intervals). However, there exist more periods where the optimized parameters clearly outperform the default parameters (morning of 2012-04-10 at 15, 10, 5, and 1-min intervals). Generally, we find that the optimized parameters provide performance one might expect based on visual classification without requiring researcher inspection of the data.

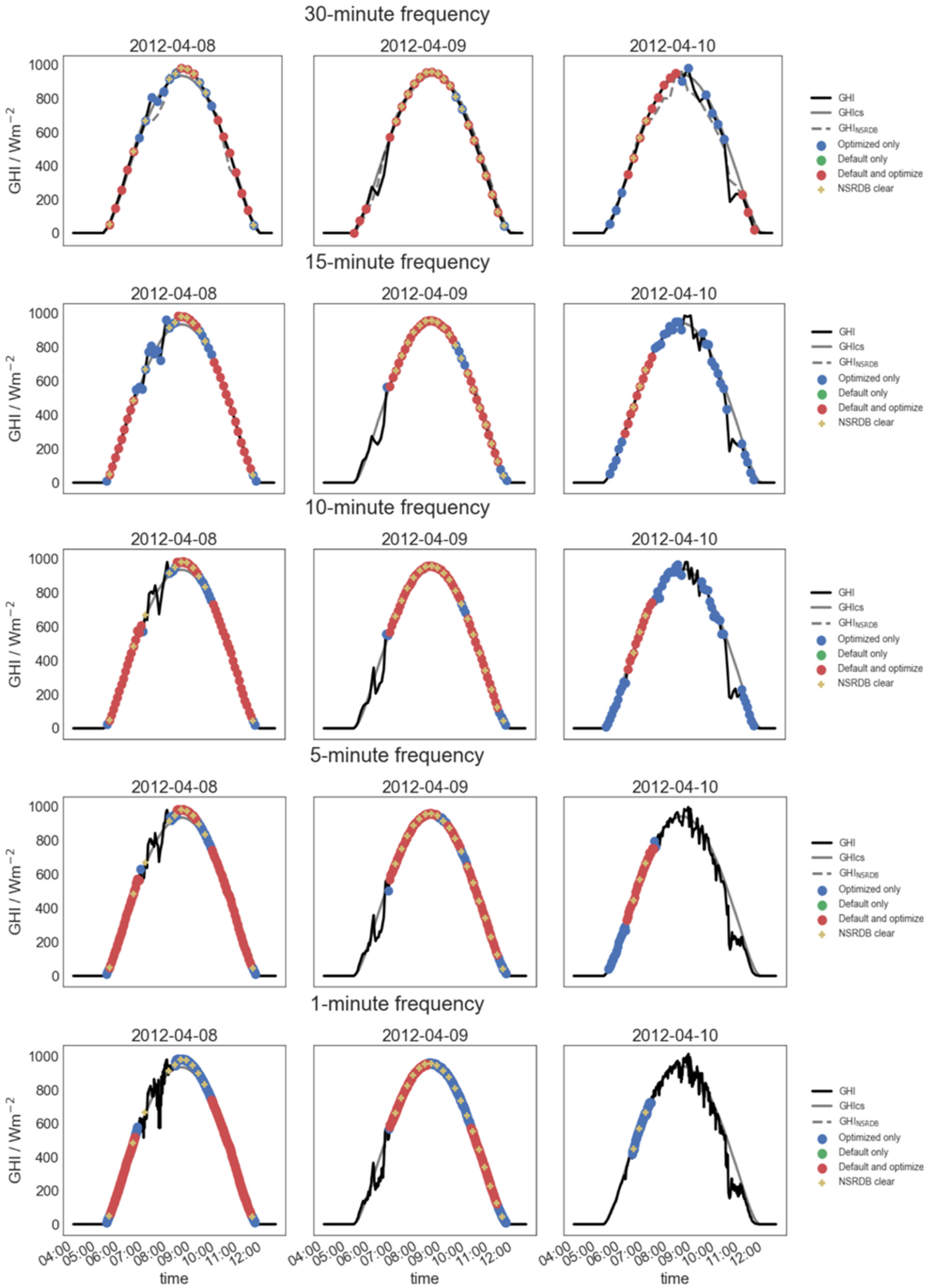

Fig. 6. Examples of ground-based GHI measurements for several days at several data intervals (black curves). Each row represents the same three days but at increasingly low intervals. The data points identified by the NSRDB database as corresponding to clear skies are labeled with yellow diamonds, whereas previous work only considered points indicated in red to be clear, the revised parameters classify both red and blue points as clear. Determinations were made using 60-min moving windows and the PVLIB parameters shown in Table II.

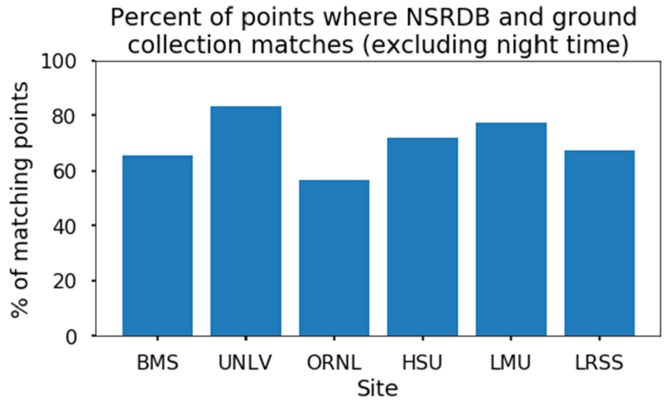

Fig. 7. Percent of data points in which ground-based and satellite-based GHI values agree to within 50 W/m^2; higher values are likely to reflect more reliable data.

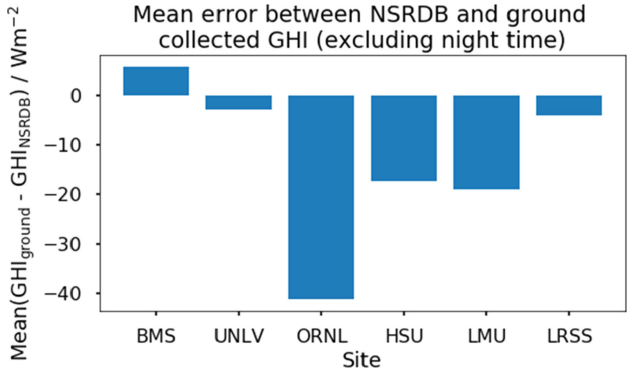

Fig. 8. Mean difference in GHI values between ground-based and satellite-based data at various sites.

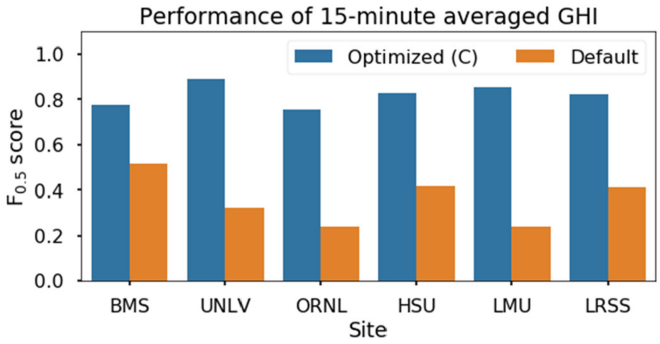

Fig. 9. $F_{0.5}$ scores for (both default and optimized algorithms) for the 15-min data interval at various sites when the GHI values are averaged (rather than instantaneously sampled) over the interval.

V. Conclusion

We developed an optimization scheme for determining parameters for PVLIB clear sky identification at six different locations across the United States with a total span of 29 years by combining ground-based irradiance measurements with satellite-based cloud cover. The parameter sets determined using this optimization scheme are capable of producing clear-sky classifications in good agreement with satellite-image-based classification while only using ground-based GHI data as input. We demonstrated that a single, universal set of optimized parameters perform consistently across multiple data intervals and geographic locations, yielding a mean $F_{0.5}$ score of 0.943 $\pm$ 0.040. This is only slightly lower than site-and-interval specific optimization, which produced an $F_{0.5}$ score of 0.949 $\pm$ .035. We expect that the findings of this paper will be of interest to researchers that require clear-sky classification routines in their PV performance and degradation analysis studies.

Appendix

A. Disagreement Between Ground and Satellite-Based Global Horizontal Irradiance Values

In cases where ground and satellite data disagree, it can be more difficult to use satellite data to train and test a clear-sky algorithm for ground-based measurements. We plot in Fig. 7 the percentage of points that are in agreement (within 50 W/m^2) for ground versus satellite measurements. We note that the ORNL site shows the poorest agreement between the two GHI values (56.6% of all day-time points).

Furthermore, we find that the magnitude and sign of disagreement is not random. In most cases, the ground-based GHI is lower than the satellite-based GHI, suggesting that many points labeled as "clear" from satellite data might be reasonably classified as "cloudy" from ground-based data. In Fig. 8, we plot the mean error between NSRDB and ground-collected GHI, showing that at the ORNL site in particular the magnitude of the disagreement is 41.3 W/m^2. These results may provide an explanation for the relatively poor classification accuracy measured for the ORNL site (as explained in the main text), as the goal of the classification is to achieve agreement between ground-based and satellite-based data.

B. Averaged Versus Instantaneous Data for Downsampling of Data Intervals

In this paper, GHI was downsampled by taking the instantaneous value of the GHI at various intervals. However, in many practical systems, GHI is downsampled by averaging over the relevant interval. We have investigated whether the $F_{0.5}$ scores for the 15-min data interval GHI change when using an averaging versus instantaneous approach (see Fig. 9). We find that the $F_{0.5}$ scores indeed drop and may need to be reoptimized for using averaged GHI. Nevertheless, the as-trained parameters as reported in our paper significantly outperform the previous parameters even with no further tuning.

Acknowledgment

The authors thank M. Sengupta and A. Habte for insightful discussions about the NSRDB data. The views expressed in the paper do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publications, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published from of this work, or allow others to do so, for U.S. Government purposes.

REFERENCES

[1] D. C. Jordan, C. Deline, S. R. Kurtz, G. M. Kimball, and M. Anderson, "Robust PV degradation methodology and application," *IEEE J. Photovolt.*, vol. 8, no. 2, pp. 525–531, Mar. 2018.

[2] J. Stein, C. Hansen, and M. Reno, "The variability index: A new and novel metric for quantifying irradiance and PV output variability," in *Proc. World Renew. Energy Forum*, 2012, pp. 1–7.

[3] M. J. Reno and C. W. Hansen, "Identification of periods of clear sky irradiance in time series of GHI measurements," *Renew. Energy*, vol. 90, pp. 520–531, 2016.

[4] S. Younes and T. Muneer, "Clear-sky classification procedures and models using a world-wide data-base," *Appl. Energy*, vol. 84, no. 6, pp. 623–645, 2007.

[5] M. Cucumo, A. De Rosa, V. Ferraro, D. Kaliakatsos, and V. Marinelli, "Correlations of direct solar luminous efficacy for all sky, clear sky and intermediate sky conditions and comparisons with experimental data of five localities," *Renew. Energy*, vol. 35, no. 10, pp. 2143–2156, 2010.

[6] C. E. Duchon and M. S. O'Malley, "Estimating cloud type from pyranometer observations," *J. Appl. Meteorol.*, vol. 38, no. 1, pp. 132–141, 1999.

[7] R. Perez, P. Ineichen, R. Seals, J. Michalsky, and R. Stewart, "Modeling daylight availability and irradiance components from direct and global irradiance," *Sol. Energy*, vol. 44, no. 5, pp. 271–289, 1990.

[8] T. Muneer, M. S. Gul, and J. Kubie, "Models for estimating solar radiation and illuminance from meteorological parameters," *J. Sol. Energy Eng.*, vol. 122, pp. 146–153, 2000.

[9] J. Calbó, J. A. González, and D. Pagès, "A method for sky-condition classification from ground-based solar radiation measurements," *J. Appl. Meteorol.*, vol. 40, no. 12, pp. 2193–2199, 2001.

[10] D. Pagès, J. Calbó, and J. A. González, "Using routine meteorological data to derive sky conditions," *Ann. Geophys.*, vol. 21, no. 3, pp. 649–654, 2003.

[11] A. Heidinger, E. Amato, M. Foster, and A. Walther, "A naive Bayesian cloud-detection scheme derived from CALIPSO and applied within PATMOS-x," *J. Appl. Meteorol. Climatol.*, vol. 51, pp. 1129–1144, 2012.

[12] A. Habte, M. Sengupta, and A. Lopez, "Evaluation of the national solar radiation database (NSRDB): 1998 – 2015," Tech. Rep. NREL/TP-5D00-67722, Nat. Renew. Energy Lab., Golden, CO, USA, 2017.

[13] M. S. Ghonima *et al.*, "A method for cloud detection and opacity classification based on ground based sky imagery," *Atmos. Meas. Technol.*, vol. 5, no. 11, pp. 2881–2892, 2012.

[14] H. Yang *et al.*, "Solar irradiance forecasting using a ground-based sky imager developed at UC San Diego," *Sol. Energy*, vol. 103, pp. 502–524, 2014.

[15] J. Calbó and J. Sabburg, "Feature extraction from whole-sky ground-based images for cloud-type recognition," *J. Atmos. Ocean. Technol.*, vol. 25, pp. 3–14, 2008.

[16] P. Ineichen and R. Perez, "A new airmass independent formulation for the linke turbidity coefficient," *Sol. Energy*, vol. 73, no. 3, pp. 151–157, 2002.

[17] W. F. Holmgren, R. W. Andrews, A. T. Lorenzo, and J. S. Stein, "PVLIB Python 2015," in *Proc. IEEE 42nd Photovolt. Spec. Conf.*, 2015, pp. 1–5.

[18] A. Andreas and T. Stoffel, "NREL solar radiation research laboratory (SRRL): Baseline measurement system (BMS); Golden, Colorado (Data)," Tech. Rep. DA-5500-56488, Nat. Renew. Energy Lab., Golden, CO, USA, 1981. [Online]. Available: http://dx.doi.org/10.5439/1052221

[19] M. Yoder and A. Andreas, "Lowry range solar station," 2008. [Online]. Available: http://dx.doi.org/10.5439/1052550

[20] A. Andreas and S. Wilcox, "Solar resource & meteorological assessment project (SOLRMAP): Rotating shadowband radiometer (RSR); Los Angeles, California (Data)," Tech. Rep. DA-5500-56502, Nat. Renew. Energy Lab., Golden, CO, USA, 2012. [Online]. Available: http://dx.doi.org/10.5439/1052230

[21] C. Maxey and A. Andreas, "Oak ridge national laboratory (ORNL); rotating shadowband radiometer (RSR); Oak Ridge, Tennessee (Data)," Tech. Rep. DA-5500-56512, Nat. Renew. Energy Lab., Golden, CO, USA, 2007. [Online]. Available: http://dx.doi.org/10.5439/1052553

[22] A. Andreas and S. Wilcox, "Solar radiation monitoring station (SoRMS): Humboldt State University, Arcata; California (Data)," Tech. Rep. DA-5500-56515, Nat. Renew. Energy Lab., Golden, CO, USA, 2007. [Online]. Available: http://dx.doi.org/10.5439/1052559

[23] A. Andreas and T. Stoffel, "University of Nevada (UNLV): Las Vegas, Nevada (Data)," Tech. Rep. DA-5500-56509, Nat. Renew. Energy Lab., Golden, CO, USA, 2006. [Online]. Available: http://dx.doi.org/10.5439/1052548